

# 证 明

本证明之附件是向本局提交的下列专利申请副本

申 请 日： 2002 11 14

申 请 号： 02 1 49375.8

申 请 类 别： 发明

REC'D 23 DEC 2003

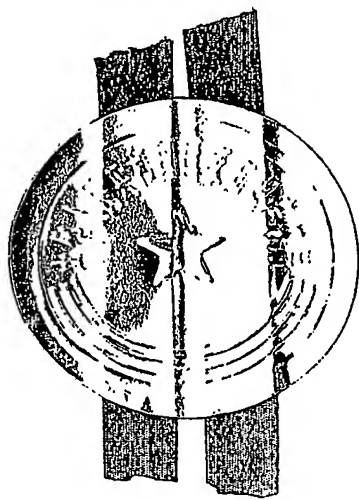
WIPO

PCT

发明创造名称： 一种编码鸡 II 型胶原蛋白的全长多核苷酸序列及其用途

申 请 人： 中国人民解放军军事医学科学院附属医院

发明人或设计人： 奚永志； 刁彩霞



**PRIORITY  
DOCUMENT**  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

中华人民共和国  
国家知识产权局局长

王景川

2003 年 11 月 28 日

## 权 利 要 求 书

1. 一种如 SEQ ID NO: 1 所示的包含编码全长鸡 II 型胶原蛋白的多核苷酸序列之分离的核酸分子, 或其具有相同生物学功能的片段。
2. 权利要求 1 所述的分离的核酸分子, 其为具有如 SEQ ID NO: 2 所示编码全长鸡 II 型胶原蛋白的多核苷酸序列片段。
3. 权利要求 1 所述分离的核酸分子或其具有相同生物学功能的片段编码的鸡 II 型胶原蛋白, 或其具有相同生物学活性的片段。
4. 一种含有权利要求 1 或 2 所述核酸分子或其具有相同生物学功能的片段的重组表达载体。
5. 一种由权利要求 4 所述重组表达载体转化的宿主细胞, 其能够表达鸡 II 型胶原蛋白, 或其具有相同生物学活性的片段。
6. 一种制备权利要求 3 所述鸡 II 型胶原蛋白的方法, 其中包括
  - 1). 用权利要求 4 所述重组表达载体转化合适的宿主细胞;
  - 2). 在合适的培养基中及适当的培养条件下培养宿主细胞;
  - 3). 从培养基或细胞中分离、纯化目的蛋白质。
7. 权利要求 6 所述方法制备的鸡 II 型胶原蛋白用于制备治疗和 / 或预防类风湿性关节炎药物的用途。
8. 一种用于预防和 / 或治疗骨关节炎、类风湿性关节炎的药物组合物, 其中包含治疗有效量的根据权利要求 6 所述方法制备的鸡 II 型胶原蛋白, 和任选的, 药学可接受的载体。
9. 一种食品或饮料组合物, 其特征在于包含一定量的权利要求 6 所述方法制备的鸡 II 型胶原蛋白。
10. 一种食品添加剂组合物, 其中包含一定量的权利要求 6 所述方法制备的鸡 II 型胶原蛋白。
11. 权利要求 1 所述核酸分子或其具有相同生物学功能的片段用于基因治疗的用途。

# 说明书

## 一种编码鸡 II 型胶原蛋白的全长多核苷酸序列及其用途

### 发明领域

本发明涉及一种包含如 SEQ ID NO: 1 所示编码全长鸡 II 型胶原蛋白的多核苷酸序列之核酸分子, 或其具有相同生物学功能的片段; 及其所编码的鸡 II 型胶原蛋白。本发明还涉及一种制备所述鸡 II 型胶原蛋白的方法, 以及由本发明所述方法制备的鸡 II 型胶原蛋白用于制备治疗和 / 或预防类风湿性关节炎的药物的用途。本发明特别涉及一种用于预防和 / 或治疗骨关节炎、类风湿性关节炎的药物组合物, 以及食品或饮料组合物或食品添加剂组合物, 其中含有本发明方法制备的所述鸡 II 型胶原蛋白。本发明还涉及所述核酸分子用于基因治疗的用途。

### 发明背景

类风湿性关节炎 (RA) 是严重影响人类健康的常见病、多发病, 由于迄今为止对其病因及其发病机制不甚了解, 因此控制炎症、缓解症状、维持关节功能仍是目前 RA 治疗的主要手段, 远未达到控制关节损伤的目标。近年来, 随着分子生物学技术的发展, 对 RA 研究取得了巨大进展, 尤其是随着对其发病机制认识的不断深入, 新的治疗方法和策略也相应产生。

1993 年, 美国科学家首次在《科学》(Trentham DE, Dynesius-Trentham RA, Orav EJ, et al. Science, 1993, 261: 1727-1730) 报道了他们用天然鸡 II 胶原蛋白 (鸡 CCII) 治疗 RA 患者获得巨大成功的研究结果, 立刻引起全世界对这一疗法的高度重视。目前美国、英国、法国等发达国家正在进行鸡 II 型胶原蛋白治疗 RA 的 II-III 期临床实验研究。从已发表的研究结果综合分析可以得出如下基本结论: 口服耐受的确为 RA 的治疗开辟了一条能标本兼治极具前景的新

途径、新策略和新疗法(Trentham DE, Dynesius-Trentham RA, Orav EJ, et al., Science, 1993,261: 1727-1730,).

目前, 在鸡 II 型胶原蛋白还未完成 II-III 期临床实验之前, 欧美发达国家制药公司纷纷将鸡 II 型胶原蛋白制成食品添加剂, 以避免冗长烦琐的药物审批。然而, 纵观目前国际上探索口服免疫耐受治疗 RA 所用的均是天然鸡 II 型胶原蛋白, 其最大的缺陷就在于: 1) 不同公司生产制备的天然鸡 II 型胶原蛋白质量不同, 即便是同一公司所生产的不同批号的鸡 II 型胶原蛋白产品质量也会不同。从而也就难以确保疗效的一致性、连续性。2) 天然鸡 II 型胶原蛋白提取和制备所需人力物力甚大。

本发明人提出了解决上述问题的新思路: 即采用基因工程方法重组生产鸡 II 型胶原蛋白。为此, 提出克隆其编码基因, 即 CCOL2A1 基因, 并实现高效表达, 重组生产鸡 II 型胶原蛋白的设想。本发明就是基于发明人首次克隆成功编码鸡 II 型胶原蛋白的全长多核苷酸序列-CCOL2A1 cDNA 而完成的。

### 发明概述

本发明一个方面涉及一种如 SEQ ID NO: 1 所示包含编码全长鸡 II 型胶原蛋白的多核苷酸序列之核酸分子, 或其具有相同生物学功能的片段。

本发明又一方面涉及由本发明所述核酸分子编码的鸡 II 型胶原蛋白, 或其具有相同生物学活性的片段。

本发明另外涉及一种含有本发明所述编码鸡 II 型胶原蛋白的核酸分子或其具有相同生物学功能的片段的重组表达载体。

本发明又一方面涉及一种由上述重组表达载体转化的宿主细胞, 其能够表达鸡 II 型胶原蛋白, 或其具有相同生物学活性的片段。

本发明再一方面涉及一种制备所述鸡 II 型胶原蛋白的方法, 其中包括

- 1). 用本发明所述重组表达载体转化合适的宿主细胞;

2).在合适的培养基中及适当的培养条件下培养宿主细胞;

3).从培养基或细胞中分离、纯化目的蛋白质。

本发明还涉及了本发明所述鸡 II 型胶原蛋白用于制备治疗和 / 或预防类风湿性关节炎的药物的用途。

本发明特别涉及了一种用于预防和 / 或治疗骨关节炎、类风湿性关节炎的药物组合物, 其中包含治疗有效量的由本发明所述方法制备的鸡 II 型胶原蛋白, 和任选的, 药学可接受的载体。

本发明还包括食品或饮料组合物, 其特征在于包含一定量的由本发明所述方法制备的鸡 II 型胶原蛋白。

本发明也包括一种食品添加剂组合物, 其中包含一定量的由本发明所述方法制备的鸡 II 型胶原蛋白。

本发明还包括了本发明所述核酸分子或其片段用于基因治疗的用途。

### 发明详述

口服耐受是自九十年代国际上兴起治疗自身免疫病的特异性免疫疗法之一, 也是近年来免疫学研究领域中最活跃、最富有成效的领域之一。它是指口服某种蛋白质抗原, 随后再用该抗原进行胃肠外免疫, 引起机体对该抗原产生全身性低免疫应答状态。现已知介导口服耐受的主要机制包括自身细胞抑制、克隆无能、克隆清除以及旁观者抑制, 这还远不是其机制的全部。其中决定因素是抗原的用量、种类来源、性质、抗原递呈处理的过程以及宿主的遗传背景和发育程度。

从已发表的研究结果综合分析可以得出如下基本结论: 1).口服耐受的确为 RA 的治疗开辟了一条能标本兼治极具前景的新途径、新策略和新疗法; 2).口服耐受鸡 II 型胶原蛋白的确对相当一部分 RA 患者疗效十分显著; 3).鸡 II 型胶原蛋白的疗效要远好与其它物种来源的 II 胶原蛋白, 因其富含大量的硫酸软骨素 A 和粘蛋白 (proteo-glycans), 而粘蛋白中硫酸葡萄糖胺含量最高, 它们有着强有力的抗炎作用和软骨修复作用; 尤其是来源于 6-8 周龄雏鸡胸骨的鸡 II

型胶原蛋白 其含量最高; 4).此外,II 胶原蛋白中还含有抗氧化作用的粘蛋白, 又称之为软骨基质糖蛋白(CMGP), 它能有效地减少氧化作用对软骨细胞的损伤; 5) 鸡 II 型胶原蛋白能有效防止蛋白酶对关节软骨的消化破坏, 重新编程已被破坏的软骨细胞和细胞因子, 从而明显减少炎症的发生; 6). 鸡 II 型胶原蛋白能促进软骨细胞及粘蛋白的合成, 增加关节滑液及透明质酸的分泌.; 7). 鸡 II 型胶原蛋白是强有力的抗炎剂和疼痛缓解剂; 8) 鸡 II 型胶原蛋白治疗十分安全无任何毒副作用, 这是目前任何其它治疗 RA 药物所不能比拟的。

目前,口服 II 型胶原诱导机体产生免疫耐受已成为有效治疗类风湿关节炎极为重要的新策略。为了提供大量、优质鸡 II 型胶原蛋白用于上述治疗, 本发明人摒弃传统的从天然来源中提取、纯化的方法, 采用基因工程方法重组生产鸡 II 型胶原蛋白。

为此, 本本发明人对鸡 II 型胶原蛋白编码基因进行了克隆、染色体定位以及特性分析。鸡 II 型胶原蛋白 cDNA 基因序列较长, 全长为 5047 bp (含外显子 2), 而目前对其功能区的 cDNA 序列所知甚微, 因此获得鸡 II 型胶原蛋白三螺旋区的 cDNA 无疑非常重要。但 II 型胶原具有复杂的二级结构, 高重复序列, 高 GC 含量 (平均 GC 含量大于 70%, 个别区域 GC 含量高达 80%) (Nah DH, Upholt WB., J Biol Chem, 1991, 266 34:23446-23452), 因此, 直接进行 PCR 非常困难。

因此, 我们首先采用 Goldkey 软件对已知的鸡 II 型胶原蛋白 3' 原肽区 cDNA 序列进行酶切位点分析, 初次尝试采用限制性内切酶不完全消化 cDNA, 再用消化后的 cDNA 进行 C 端原肽区的 PCR 扩增并获成功。鉴于鸡 II 型胶原蛋白基因序列较长, 我们将 cDNA 分成五部分并进行分部 PCR, 各扩增片断彼此之间重叠至少 50bp, 便于 SOE-PCR 连接。针对胶原基因中的高 GC 含量, PCR 扩增中选用适宜扩增高 GC 含量且保证性强的 Tag 酶。

胶原在机体内广泛存在, 不同的胶原型可公用其中的一条链, XI 型胶原与 II 型胶原之间则属于此情况。XI 型胶原是由三个不同的多肽亚单位  $\alpha 1$ 、 $\alpha 2$ 、 $\alpha 3$  组成的, 其中  $\alpha 3$  是  $\alpha 1$  (II) 基因的表达产物, 即

7

$\alpha 3$  与 II 型胶原  $\alpha 1$  链仅羟化赖氨酸含量不同(Rousseau JC, Farjanel J, Boutillon MM, et al., J Biol Chem, 1996,271(39): 23743-8); XIII 型胶原为 II 型胶原的跨膜形式, 即 XIII 型胶原与 II 型胶原链的组成完全相同(Snellman A,Keranen MR, Hagg PO, et al. J Biol Chem, 2000, 275(12):8936-44).

我们在成年鸡 II 型胶原蛋白表达谱的检测中, 由于 ELISA 实验中鸡 II 型胶原蛋白经胃蛋白酶、弹性蛋白酶消化后, II 型胶原形成  $\alpha$  链小短肽, 因此含有  $\alpha 1$  (II) 链的胶原均可与鸡 II 型胶原蛋白单克隆抗体结合, 因此实验中不能排出同时检测到 XI, XIII 的可能。

II 型胶原外显子 2 在不同的组织采取不同的剪接方式。在实验中我们发现 17 日龄鸡胚胸骨、生长板、关节软骨中获取的鸡 II 型胶原蛋白 cDNA 不含外显子 2; 而肝脏、小肠、肌肉、皮肤中以含有外显子 2 的产物占多数; 其中眼球玻璃体、角膜中提取的总 RNA, 经 RT-PCR 发现存在含与不含有外显子 2 两种情形; 但眼球玻璃体 RT-PCR 显示, 在较小日龄的鸡胚 (小于 14 天) 仅以含外显子 2 的形式存在 (数据未给出); 而当鸡胚发育至 17 天, RT-PCR 结果却出现另外一种情形, 以不含外显子 2 为主要表达形式。上述现象只能以下面的解释作为答案: 鸡 II 型胶原蛋白在软骨组织中已不含外显子 2 的形式存在, 而在非软骨组织中如心脏、肝脏、角膜、小肠、肌肉、皮肤等中主要以含有外显子 2 的形式存在; II 型胶原基因组存在调控序列, 其调控 II 型胶原在胚胎发育的不同时期及不同组织中的表达; 鸡 II 型胶原蛋白在各脏器广泛存在的现象随着鸡的成长发育而消失, 而代之以鸡 II 型胶原蛋白仅在软骨组织中表达; 发育的鸡胚胸骨中鸡 II 型胶原蛋白含量丰富 (Young MF, Vogeli G, Nunez AM, et al., Nucleic Acids Res, 1984,12 (10): 4207-4228; Marshall GE, Konstas AGP, Lee WR., BrJ Ophthalmol, 1993,77:515-524), 但随鸡的生长发育, 胸软骨逐渐钙化, 鸡 II 型胶原蛋白含量逐渐下降。

关于 II 型胶原在机体内的表达, 目前认为 II 型胶原在成熟的个体存在于眼球玻璃体与软骨 (Seery CM, Davison PF., Invest

Ophthalmol Vis Sci,1991,32:1540-1550; Huerre-Jeanpierre C, Mattei MG, Weil D, et al., Am J Hum Genet, 1986, 38(1): 26-37), 而我们对天然 II 型胶原的表达检测中却发现眼球玻璃体中不存在 II 型胶原的表达。

鸡 II 型胶原蛋白基因组 DNA 具有复杂的结构, 我们通过 PCR 对鸡外周血 (EDTA 抗凝) 中提取的 DNA 进行扩增, 仅能获得鸡 II 型胶原蛋白 3' 端 5494bp 的片段, 而 3' 端以外的基因组序列, 无论模板、引物、PCR 扩增条件如何调整, 均不能得到目的基因, 究其原因, 可能是由于鸡 II 型胶原蛋白基因含量高 GC, 重复序列多, 引物非特异结合部位太多而造成 PCR 扩增的失败。鸡 II 型胶原蛋白 3' 端基因组 DNA 测序结果可看出, II 型胶原在进化上高度保守, 不同种属的 II 型胶原蛋白其氨基酸同源性均高于 92%, 甚至达 99%。鸡 II 型胶原蛋白与 I 型胶原、III 型胶原蛋白相比, 其内含子外显子明显的小, 这与 Upholt 等曾报道的相同 (Ausar SF, Beltramo DM, Castagna LF, et al., Rheumatol Int, 2001, 20:138-144)。

RH 杂交作图是目前基因在染色体定位最精确的方法, 但限于鸡 RH 杂交板目前尚无厂家出售。因此我们选用 ISH 杂交来进行定位, 鸡 II 型胶原蛋白编码基因定位于 4 号染色体, 而小鼠、人相应的 COL2A1 则分别定位于 8 号及 12 号染色体 (Barnett ML, Combitchi D, Trentham DE., Arthritis & Rheumatism, 1996, 39 4:623-628)。

为实现本发明的目的, 本发明人对鸡 II 型胶原蛋白编码 cDNA 成功进行了克隆、染色体定位、表达谱分析及基因组 DNA 克隆。尤其鸡 II 型胶原蛋白 cDNA 的克隆成为本发明完成的基础, 意义重大。鸡 II 型胶原蛋白作为口服耐受治疗 RA 的新策略, 其 cDNA 的克隆无疑使基因工程方法生产鸡 II 型胶原蛋白成为可能。

具体的, 本发明一个方面涉及一种如 SEQ ID NO: 1 所示包含编码全长鸡 II 型胶原蛋白的多核苷酸序列之分离的核酸分子, 或其具有相同生物学功能的片段。

本发明又一方面涉及由本发明所述分离的核酸分子编码序列所



编码的鸡 II 型胶原蛋白，或其具有相同生物学活性的片段或其保守性变体。

本发明中，蛋白质或多核苷酸“变体”是指一种具有一个或多个氨基酸或核苷酸改变的氨基酸序列或编码它的多核苷酸序列。所述改变可包括氨基酸序列或核苷酸序列中氨基酸或核苷酸的缺失、插入或替换。变体可具有“保守性”改变，其中替换的氨基酸具有与原氨基酸相类似的结构或化学性质，如用亮氨酸替换异亮氨酸。变体也可具有非保守性改变，如用色氨酸替换甘氨酸。

“缺失”是指在氨基酸序列或核苷酸序列中一个或多个氨基酸或核苷酸的缺失。

“插入”或“添加”是指在氨基酸序列或核苷酸序列中的改变导致与天然存在的分子相比，一个或多个氨基酸或核苷酸的增加。

“替换”是指由不同的氨基酸或核苷酸替换一个或多个氨基酸或核苷酸。

“生物活性”是指具有天然分子的结构、调控或生物化学功能的蛋白质。

“相似性”是指氨基酸序列之间排列对比时相应位置氨基酸残基的相同或保守性取代的程度。用于保守性取代的氨基酸例如，带负电荷的氨基酸可包括天冬氨酸和谷氨酸；带正电荷的氨基酸可包括赖氨酸和精氨酸；具有不带电荷的头部基团有相似亲水性的氨基酸可包括亮氨酸、异亮氨酸和缬氨酸；甘氨酸和丙氨酸；天冬酰胺和谷氨酰胺；丝氨酸和苏氨酸；苯丙氨酸和酪氨酸。

“分离的”一词指将物质从它原来的环境（例如，若是自然产生的就指其天然环境）之中移出。比如说，一个自然产生的多核苷酸或多肽存在于活动物中就是没有被分离出来，但同样的多核苷酸或多肽同一些或全部在自然系统中与之共存的物质分开就是分离的。这样的多核苷酸可能是某一载体的一部分，也可能这样的多核苷酸或多肽是某一组合物的一部分。既然载体或组合物不是它天然环境的成分，它们仍然是分离的。

本发明的多核苷酸可以RNA或DNA的形式存在，其中DNA包括cDNA、基因组DNA和合成的DNA。DNA可以是双链或者单链的，单链也可以是有义链或者反义链。编码多肽的多核苷酸序列可以与SEQ ID No.1所示的多核苷酸序列相同；或者可以是一个由于遗传密码的冗余或简并性而不同的多核苷酸序列，但它编码与SEQ ID No.2相同的成熟多肽。

本发明进一步涉及利用本发明所述的多核苷酸的各种变体，它们编码含有SEQ ID No.2所示的推定的氨基酸序列多肽的片段、类似物和衍生物。这些多核苷酸变体可以是天然存在的等位基因变体或非天然存在的多核苷酸变体。

编码本发明多肽的多核苷酸也可能与特定的标记序列在同一读框中相融合，标记序列可帮助本发明多肽的纯化。标记序列在使用细菌宿主时可以是pQE载体的六聚组氨酸标记，以利于融合有标记的成熟多肽的纯化，或者当使用哺乳类宿主（如猴肾成纤维细胞的COS-7细胞系）时，标记序列可以是一种血细胞凝集素（HA）标记。此外，包含编码本发明多肽的多核苷酸序列还可包含同源或异源的特定信号肽序列，以帮助目的蛋白质分泌到原核细胞或真核细胞膜外。本领域普通技术人员知晓，上述标记序列和信号肽序列也可通过重组方法或化学法添加在用于表达本发明多肽的载体上。

本发明进一步涉及具有SEQ ID No.2所示推定氨基酸序列的多肽及其活性片段、类似物和衍生物。

SEQ ID No.2所示多肽的“具有相同生物学活性的片段”指能基本保留该多肽的生物学功能或活性的多肽。

由本发明所述方法制备的鸡II型胶原蛋白为重组多肽或其片段、衍生物和类似物。具体的，SEQ ID No.2所示多肽的片段、衍生物和类似物可以是：(i)一个或多个氨基酸残基被保守性或非保守性氨基酸残基所取代（优选是保守性氨基酸残基）的多肽，取代的氨基酸残基是或不是由遗传密码所编码的氨基酸。例如，可以通过氨基酸残基的插入、取代和/或删除，得到鸡II型胶原蛋白的沉默突变体或功能等同物。

13

可基于氨基酸残基之间在极性、电荷、溶解性、疏水性、亲水性和/或两亲性等方面的相似性进行保守性氨基酸取代，只要保留鸡II型胶原蛋白的活性；或(ii)一个或多个氨基酸残基带有取代基团的多肽；或(iii)成熟多肽与其它功能性化合物，如提高多肽半寿期的化合物（例如聚乙二醇）融合在一起的多肽；或(iv)成熟多肽与其它氨基酸序列相融合的多肽，其中所述其它氨基酸序列包括诸如帮助纯化成熟蛋白的氨基酸序列或蛋白原序列等。这些帮助纯化的结构域包括但不限于：金属螯合肽，如用于在固定化金属上纯化的组氨酸-色氨酸模块；用于在固定化免疫球蛋白上纯化的蛋白A结构域以及用于FLAGS延伸/亲和纯化系统的结构域（IMMUNEX公司，Seattle, Wash.）。特异于因子XA或肠激酶的断裂接头序列也可用于帮助目的蛋白质的纯化（Porath, J. 等人(1992), Prot.Exp.Purif. 3:263-281）。从这些公开内容看，这样的片段、衍生物和类似物的应处于本领域技术人员的知识范围内。

本发明的鸡II型胶原蛋白包括SEQ ID No.2所示的鸡II型胶原蛋白，即成熟多肽，也包括与SEQ ID No.2多肽至少有至少有90%相似性（优选是90%的相同性）的多肽，更优选至少有95%相似性（优选是95%相同性）的多肽，也包括这些多肽的一些部分，通常这些多肽部分至少含有30个氨基酸、优选至少50个氨基酸。

本发明的多肽、其保守性变体和生物活性片段及衍生物可以用常规的肽合成的方法制备，例如固相肽合成（Merrifield J. (1963), 美国化学学会杂志 (J. Am. Chem. Soc.) 85:2149-2154; Roberge, J. Y. 等人, (1995) 科学, 269:202-204)。蛋白质合成可以手工完成，也可利用肽自动合成仪如Applied Biosystems 431A 肽合成仪进行（Perkin Elmer）。本发明鸡II型胶原蛋白也可以从天然生物材料中经分离纯化得到，但优选利用本发明提供的多核苷酸用重组DNA技术制备。

根据常规的重组DNA技术，利用本发明的多核苷酸序列可表达或制备重组的鸡II型胶原蛋白。本发明因而涉及制备本发明鸡II型胶原蛋白的方法，一般包括以下步骤：

(1).用本发明编码鸡 II 型胶原蛋白的多核苷酸（或变体）或含有该多核苷酸的重组表达载体转化合适的宿主细胞；

(2).在合适的培养基中及适当的培养条件下培养宿主细胞；

(3).从培养基或细胞中分离、纯化目的蛋白质。

本发明也涉及包含本发明多核苷酸的重组载体、带有本发明重组载体的遗传工程化宿主细胞和通过重组技术制备本发明多肽的方法。

本发明的多核苷酸可以用来经重组技术产生多肽。例如，该多核苷酸可以存在于选自多种用于表达多肽的表达载体中的任一载体上，这些载体包括染色体的、非染色体的及合成的DNA序列，例如，SV40的衍生物、细菌质粒、噬菌体DNA、酵母质粒、衍生于质粒和噬菌体DNA结合的载体、病毒DNA、杆状病毒、牛痘病毒、腺病毒、禽痘病毒及伪狂犬病毒，包括但不限于pQE系列(Qiagen)、pBS、pD10、phagescript、psiX174、pBluescript SK、pBSKS、pNH系列(Stratagene)、pTRC99a、pKK223-3、pDR540、pRIT5 (Pharmacia); pWLNEO、pSV2CAT、pOG44、pXT1、pSG (Stratagene)、pSVK3、pBPV、pMSG、pSVL (Pharmacia)。不过，只要是能在宿主中复制和存活，其它质粒或载体也可应用。

本发明也包括含有本发明多核苷酸的重组构建体。该构建体包括载体，如上述质粒或病毒载体，其中可正向或反向插入本发明的多核苷酸序列。构建体中还包含调节序列，例如，有效地连接到本发明多核苷酸序列上的启动子（包括组成型和诱导型启动子），介导下游结构序列的转录。合适的启动子包括但不限于， $\lambda$ 噬菌体的PL启动子、杆状病毒多角体蛋白启动子；细菌启动子如：LacI、LacZ、T3、T7、gpt、 $\lambda$  PR、PL和trp；真核启动子如：CMV立即早期启动子、HSV胸苷激酶启动子、早期和晚期SV40启动子、逆转录病毒的LTR和小鼠金属硫蛋白I启动子；还包括衍生自植物细胞基因组中的启动子，如热休克蛋白、RUBISCO启动子。

表达载体也包含翻译起始所需要的核糖体结合位点和转录终止子，其还可以具有增强表达的合适序列，如增强子。增强子是DNA的

顺式作用因子，通常有大约10-300bp，作用于启动子，提高它的转录。例如，SV40中位于复制起点后侧100-270bp处的增强子，巨细胞病毒早期启动子增强子、位于复制起点后侧的多形瘤增强子及腺病毒增强子。此外，表达载体优选还包含能提供表型特征的一种或多种选择性基因、抗性基因和/或标记基因，以便于转化宿主细胞的筛选。选择性基因如帮助细胞利用吲哚或组氨酸的trpB、hisD (Hartman, S. C., 和R.C.Mulligan(1988) 美国国家科学院院报 85:8047-51), 用于tk<sup>-</sup>或aprt<sup>-</sup>细胞的肝疱疹病毒胸腺嘧啶激酶 (Wigler, M. 等人(1977) 细胞 11:223-32) 和腺嘌呤磷酸核糖转移酶 (Lowy, I. 等人 (1980) 细胞 22:817-23); 抗性基因如赋予氨苄唑啉抗性的二氢叶酸还原酶DHFR (Wigler, M. 等人(1989), 美国国家科学院院报, 77:3567-70) 或赋予新霉素和G-418抗性的npt (Colbere-Garapin, F.等人(1981) 生物化学杂志 (J.Mol.Biol.), 150:1-14), 以及四环素或氨苄青霉素抗性基因。哺乳类表达载体一般包括复制起点、适当的启动子、增强子及任何必需的核糖体结合位点、多腺苷化位点、拼接供体和受体位点、转录终止序列和5'侧翼非转录序列。衍生于SV40拼接序列的DNA序列和多腺苷化位点可用于提供所需要的非转录遗传元件。此外，包含编码本发明多肽的核苷酸序列的载体还可包含同源或异源的特定信号肽序列，以帮助目的蛋白质分泌到原核细胞或真核细胞膜外。本领域普通技术人员知晓，上述表达载体及构建体中含有的标记序列和信号肽序列也可通过重组方法或化学法添加在编码本发明多肽的多核苷酸上。

合适载体和启动子的选择为本领域普通技术人员周知。细菌适用的有效表达载体可以这样来构建：将编码目的蛋白的结构DNA序列随同适当的翻译起始和终止信号被插入到带有一个功能启动子的可操纵的阅读框中。本领域普通技术人员周知用于构建含有本发明核苷酸序列以及合适的转录及翻译调控元件的方法。这些方法包括体外重组DNA技术、合成技术以及体内遗传重组技术 (Sambrook, J. (1989), 分子克隆实验室手册, Cold Spring Harbor Press; Plainview, N.Y.; Ausubel, F.M. (1989) 当代分子生物学方法 (Current Protocols

in Molecular Biology), John Wiley & Sons, N.Y.).

包含上述合适的DNA序列、合适的启动子或控制序列的载体可以用于转化合适的宿主,让宿主表达该蛋白。

本领域技术人员知晓,根据本发明DNA序列所插入的表达载体或构建体的种类和特性选择合适的宿主以表达目的蛋白质。适于表达本发明的多肽的宿主包括但不限于:原核宿主,诸如大肠杆菌、芽孢杆菌属、链霉菌属等;真核宿主,诸如:酵母属、曲霉属、昆虫细胞,诸如果蝇S2和草地夜蛾Sf9;动物细胞,如CHO、COS(猴肾成纤维细胞系, Gluzman (细胞 23:175, 1981)及其它的能表达相容载体的细胞系,例如C127、3T3、CHO、HeLa、BHK、Bowes黑素瘤细胞;植物细胞以及腺病毒等等。各种哺乳动物细胞的培养系统也能用于表达重组蛋白。从这些讲授看,合适宿主的选择应该在本领域技术人员的知识范围内。

带有如上所述的含有本发明核苷酸序列之载体或构建体能够通过传统的方法导入合适的宿主细胞中以产生重组产物。将构建体导入上述宿主细胞的方法为本领域技术人员周知,包括但不限于:氯化钙介导的转化、磷酸钙转染、DEAE-葡聚糖介导的转染、电穿孔、显微注射、粒子轰击法或基因枪方法(Sambrook, J.(1989), 分子克隆实验室手册, Cold Spring Harbor Press; Plainview, N.Y.; Ausubel, F.M. (1989) 当代分子生物学方法, John Wiley & Sons, N.Y.; Hobbs, S. 等人, McGraw Hill Yearbook of Science and Technology (1992), McGraw Hill, N.Y.191-196; Engelhard, E.K.等人, 美国国家科学院院报, 91:3224-3227; Logan, J.等人, 美国国家科学院院报, 81:3655-3659)。

在适当的培养条件与培养基中培养经转化的宿主菌株或细胞,使其生长到恰当的细胞密度之后,用适当的方法(例如温度转变或化学药品诱导)诱导所选择的启动子,并将细胞再培养一段时间。针对不同的宿主菌株或细胞选择以及所表达的目的蛋白质的性质相应的培养条件和培养基在本领域技术人员知识范围之内。

在合适的启动子控制下可以在哺乳类细胞、酵母、细菌或其它细胞中表达成熟蛋白。利用由本发明的DNA构建体衍生的RNA,也可以用无细胞翻译体系产生这种蛋白质(Sambrook, J. (1989), 分子克隆实验室手册, 第18章第4节, Cold Spring Harbor Press; Plainview, N.Y. )。

通常用离心的方法收获细胞或培养液。对于目的蛋白质保留在胞内的情况,一般可用任何便捷的物理、化学方法或酶法,包括冻融循环、超声波、机械破碎,或使用细胞溶解剂或特定的酶破碎细胞,将所得粗提物保留以进一步纯化。对于对于目的蛋白质分泌到胞外的情况,可直接从培养上清液中利用常规方法回收目的蛋白质。这些方法都是本领域的技术人员熟知的。

将多肽从重组细胞培养物中回收和纯化的方法有硫酸铵或乙醇沉淀、酸提取、阴离子或阳离子交换层析、体积排阻层析、疏水相互作用层析、亲和层析、羟基磷灰石层析和植物凝集素层析。形成成熟蛋白的完整构象还需要蛋白质的重折叠步骤。通常高效液相层析(HPLC)或毛细管电泳可应用于最后的纯化步骤。

本发明还涉及了根据本发明所述多核苷酸序列制备的鸡 II 型胶原蛋白在制备治疗和/或预防类风湿性关节炎等病的药物的用途。具体的,本发明所述核酸分子的编码区或其部分,如全长序列、部分序列或其突变体均能用于有效表达本发明所述鸡 II 型胶原蛋白或其功能性片段。

本发明特别涉及了一种用于预防和/或治疗骨关节炎、类风湿性关节炎的药物组合物,其中包含治疗有效量的由本发明所述方法制备的鸡 II 型胶原蛋白,和任选的,药学可接受的载体。

本发明还包括食品或饮料组合物,其特征在于包含一定量的由本发明所述方法制备的鸡 II 型胶原蛋白。在发明的一个实施方案中,所述鸡 II 型胶原蛋白用于制备保健食品或食品添加剂等产品以对尤其是对骨关节病中的类风湿性关节炎、骨关节炎等病起到保健作用。

本发明也包括一种食品添加剂组合物,其中包含一定量的由本发

明所述方法制备的鸡 II 型胶原蛋白。

本发明还包括了本发明所述核酸分子或其片段用于基因治疗的用途。

#### 附图说明

图 1。显示本发明所述鸡 II 型胶原蛋白编码基因的 PCR 克隆策略。

图 2。显示本发明所述全长鸡 II 型胶原蛋白的结构。

图 3。显示鸡胚中 II 型胶原蛋白编码基因 3' 非翻译区 (UTR) 的 PCR 结果。

图 4。显示成鸡中 II 型胶原蛋白编码基因的组织特异性表达。

图 5。II 型胶原蛋白浓度与吸光度标准曲线

图 6。鸡 II 型胶原蛋白编码基因外显子的剪切分析

图 7。鸡 II 型胶原蛋白编码基因的染色体分带。

图 8。鸡染色体中期分裂相

图 9。ISH 杂交结果。

图 10。人、犬、鼠、鸡、斑点鱼之 II 型胶原蛋白与基因同源性比较进化树。

本发明将在下面的实施例中进一步描述。但是本发明并不局限于这些实施例。

其中，分子生物学、生物化学、免疫学等常规操作均按照“分子克隆：实验指南”第 2 版中的描述进行。所述酶切条件和缓冲液均按照生产厂家的说明书操作。通常所用的温育时间大约是 37℃ 一个小时，但根据厂家的指示可以有所改变。

#### 实施例 1

##### 鸡胚胸骨总 RNA 提取

17 日龄 SPF 鸡胚 (中国农业科学院畜牧研究所, 北京) 经消毒



后破壳，取鸡胚，置冰冷生理盐水中冲洗，无菌取胸骨，迅速置 Trizol (Invitrogen) 中快速研磨，转移至 1.5ml eppendorf 管，室温放置 10min，4℃、10000g 离心 10min，将上清转移至一 DEPC 处理的 eppendorf 管，每管加 200μl 氯仿，混匀，室温放置 10mins，离心同上。将水相转移至一新 eppendorf 管，加等量异丙醇，室温再放置 10min，离心，75%乙醇洗涤沉淀，无 RNase 水溶解，-80℃保存待反转录。

### cDNA 合成

提取的总 RNA，以 3'RACE 试剂盒 (TAKARA) 中的 oligodT-3sites adaptor primer 作为下游引物进行反转录。反转录完毕 95℃处理 5 分钟灭活反转录酶并将 cDNA 置-20℃保存备用。

### 鸡 COL2A1 cDNA C-端原肽 cDNA 的 PCR 扩增

根据已知鸡 COL2A1 C 端 UTR 的序列 (Sandell LJ, Prentice HL, Kravis D, Upholt WB., J Biol Chem, 1984,259 (12) 7826-7834)，利用引物 col2a1-2F, col2a1-2R (如表 1 所示)，扩增鸡 COL2A1 C-端原肽部分。C-端原肽(COL2A1-2) cDNA 进行以下 PCR 扩增程序：引物 col2a1-2F, col2a1-2R, 10 × PCR 缓冲液; 5 μL, MgCl<sub>2</sub> (25Mm) 5 μL, dNTP Mix (10Mm) 1 μL, 引物各自为 1 μL (20pmol/μL) 1 μL, 100%甘油 2.5 μL (终浓度 5%), cDNA 2 μL, Taq 0.8 μL, 总体积 50 μL。PCR 所用程序 96℃ 10min, 96℃ 1min, 62℃ 1 min, 72℃ 3min, 72℃ 7min。PCR 反应体系内含终浓度 5%甘油，1%琼脂糖电泳检查扩增结果，并采用凝胶回收试剂盒 (Invitrogen) 回收目的基因，连接 pGEM-T 载体 (Promega)，转化 DH5α，酶切鉴定后测序。

### 鸡 COL2A1 3'UTR cDNA 的克隆

根据已知鸡 COL2A1 基因序列设计扩增 COL2A1 3'UTR 的上游

引物 col2a1-1F (参见表 1) , 采用 3' RACE 策略, 以反转录相对应的 3 sites adaptor primer ( col2a1-1R) 扩增 3'UTR, 其中, 采用 PCR 缓冲液, 引物 col2a1-1R 为 3sites adaptor primers, 采用 96℃ 5min, 96℃ 1min, 63℃ 1 min, 72℃ 30 sec ( 4 个循环), 96℃ 30sec, 61℃ 30sec, 72℃ 20sec ( 26 个循环), 72℃ 7min. 回收及转化均同前述步骤。

### 鸡 COL2A1 N-原肽 cDNA 的克隆

col2a1-5F , col2a1-5R 为上下游引物扩增 5'N 原肽。采用 PCR 缓冲液, 退火与延伸在同一温度点上。96℃ 5min, 96℃ 1min, 72℃ 50 sec ( 30 个循环), 72℃ 7min.

### 重组质粒载体的构建及序列测定

质粒提取、酶切、回收、连接、转化、构建重组质粒的鉴定等均按《分子克隆实验指南》( Sambook J, Fritsch EF, Maniatis T. Molecular cloning:a laboratory manual. 2<sup>nd</sup> ed Cold Spring Harbor Laboratory Press,1989.) 进行, 将上述构建好的 COL2A1 cDNA 系列片段基因以及 SOE 连接的 COL2A1 cDNA 系列片段及全长 COL2A1 分别插入 pGEM-T 载体 ( Promega), 酶切鉴定后测序。

其中所述质粒的构建和制备如下进行:

凝胶快速提取试剂盒回收过程: PCR 扩增产物经 1% 琼脂糖电泳, 切胶回收目的基因, 再经凝胶快速提取试剂盒回收。过程如下:

- 1) 含目的基因的凝胶置于 eppendorf 管中, 加入 1ml 溶液 L1, 50℃ 水浴 15 分钟, 使凝胶彻底溶解;
- 2) 将溶解好的凝胶加入离心柱中, 12000rpm 离心, 柱内再加入 500μl L1, 室温放置 1 分钟, 离心同上;
- 3) 加 L2 700μl, 室温放置 5 分钟, 离心同上;
- 4) 12000rpm 再离心 2 分钟, 使酒精挥发;
- 5) 将柱转移至 1 新 eppendorf 管中, 加入热( 约 70℃ )的 TE 30μl,

室温放置 2 分钟, 12000rpm 离心, 柱下的液体即为回收的目的 DNA。回收 DNA 经 1% 琼脂糖电泳, 确定浓度为 40 $\mu$ g/ $\mu$ l。

目的 DNA 与 pGEMT 载体的连接: 连接体系如下: 2 $\times$ 连接缓冲液 5 $\mu$ l, 目的 DNA 3 $\mu$ l, T 载体 1 $\mu$ l, T4DNA 连接酶 1 $\mu$ l, 共计 10 $\mu$ l, 混匀, 4 $^{\circ}$ C 放置过夜。

大肠杆菌 DH5 $\alpha$  感受态的制备参照 Sambrook J, Fritsch EF, Maniatis T. Molecular cloning: a laboratory manual. 2<sup>nd</sup> ed Cold Spring Harbor Laboratory Press, 1989. 中所述方法。

PCR 连接产物转化 DH5 $\alpha$  感受态: 取如上述连接好的 pGEMT-col2a1-1 (2, 3, 4, 5 等) 4 $\mu$ l, 加入 100 $\mu$ l 新鲜制备的感受态, 冰浴 30 分钟, 42 $^{\circ}$ C 热休克 90 秒, 再迅速置冰水中 2-3 分钟, 加入 900 $\mu$ l 无氨苄 LB, 37 $^{\circ}$ C、160rpm 轻摇 45 分钟, 加入 1MIPTG 4 $\mu$ l, 16 $\mu$ l (50mg/ml) X-gal, 涂布氨苄 LB 平板, 37 $^{\circ}$ C 下培养 16 小时至菌落大小合适为止。

阳性转化子的筛选: 挑取 IPTGX-gal 氨苄青霉素 LB 平板上的白色克隆接种含有 200 $\mu$ g/ml 氨苄青霉素的 LB 中, 37 $^{\circ}$ C, 200rpm 摇床培养 12 小时, 提取质粒。

由上述方法制备得到本发明所述 pGEMT-col2a1-2, pGEMT-col2a1-1、pGEMT-col2a1-3, pGEMT-col2a1-(1+2) 和 col2a1-5、col2a1-4 质粒

质粒采用 Not 与 Nco 双酶切鉴定阳性克隆。Not 与 Nco 双酶切体系: 10 $\times$  1 $\mu$ l, 0.1%BSA 1 $\mu$ l, 采用 Promega A7500 试剂盒提取质粒 8 $\mu$ l, Not 1 $\mu$ l, Nco 1 $\mu$ l, 水 8 $\mu$ l, 共计 20 $\mu$ l, 37 $^{\circ}$ C 酶切 2 小时。酶切完毕, 1% 琼脂糖电泳鉴定阳性克隆, 将酶切阳性的克隆送样测序。核酸测序仪为 PE 公司的 ABI377。

### 鸡 COL2A1 全长 cDNA 的 SOE-PCR

为获得 COL2A1 全长 cDNA, 本研究采用重叠延伸-PCR

(SOE-PCR)策略来进行构建(Horton RM, Hunt HD, Ho SN, et al., Gene, 1989, 77:61-68), 这需进行四次 PCR 反应。

第一次 PCR 反应是以质粒 pGEMT-col2a1-2, pGEMT-col2a1-1 PCR 产物为模板, 以 col2a-2F、col2a1-1R 为上、下游引物, 使 col2a1-2、col2a1-1 连接起来形成产物 col2a1-(2+1); 采用 Pfu DNA 聚合酶, PCR 扩增条件: 96℃ 5min, 96℃ 1min, 60℃ 1 min, 72℃ 150sec (4 个循环), 96℃ 30sec, 58℃ 30sec, 72℃ 130sec (26 个循环), 72℃ 7min.;

第二次 PCR 反应是以质粒 pGEMT-col2a1-3, pGEMT-col2a1-(1+2) PCR 产物为模板, 以 col2a-3F、col2a1-1R 为上、下游引物。PCR 扩增条件: 96℃ 5min, 96℃ 1min, 64℃ 1 min, 72℃ 120 sec (4 个循环), 96℃ 30sec, 62℃ 1min, 72℃ 90sec (26 个循环), 72℃ 7min.

第三次 PCR 反应中, 以 col2a-5F、col2a1-4R 为上、下游引物, 将 col2a1-5、col2a1-4 质粒 PCR 产物为模板。PCR 扩增条件: 96℃ 5min, 96℃ 1min, 62℃ 2 min, 72℃ 120 sec (4 个循环), 96℃ 1min, 60℃ 1min, 72℃ 90sec (26 个循环), 72℃ 7min.

第四次 PCR 反应则以 col2a-5F、col2a1-1R 为上、下游引物, 以前三次连接好的 PCR 产物 (col2a1-5+4, col2a1 1+2+3) 为模板, 可获得全长 col2a1 cDNA。PCR 扩增条件: 96℃ 5min, 96℃ 1min, 64℃ 2 min, 72℃ 210 sec (4cycles), 96℃ 1min, 62℃ 1min, 72℃ 180sec (26 cycles), 72℃ 7min.

cDNA 克隆所用引物序列如表 1 示:

表 1. 鸡 COL2A1 全长 cDNA 扩增引物序列

---

col2a1-1F	5'-	TCT ATC GCG CAC CCG TTG TGC -3'
col2a1-1R	5'-	GTC TTG TAG TGC TAC GGC TTG C -3'
col2a-12F	5'-	TTG CAG ATG TCT CCA ATA CCA G

---

-3'

col2a1-2R	5'- GCA CAA CGG CTC GGG CAA TGT GCT AAC G -3'
col2a1-3F	5'- GCT CGG AAG CAA CGG CCT CG -3'
col2a1-3R	5'- CTC GTC CCG GAC GCG ACG G -3'
col2a1-4F	5'- CGC TGC GAT CGT CAT GCG G -3'
col2a1-4R	5'- GTA GTG ACC CTA CGC CCG AG -3'
col2a1-5F	5'- ACG CCG GCT CTC GTG CTC CTC GTG GTG C -3'
col2a1-5R	5'- CCG CCC GGG TCC GAA TGC CCG CAT -3'

本研究采用重叠延伸-PCR (SOE-PCR) 策略来获得全长 4837bp COL2A1 cDNA (图 1)。由于鸡 COL2A1 基因 GC 含量较高, 基因扩增过程中反应体系采用适宜扩增高 GC 含量的缓冲液, 扩增产物特异性极高, 将其插入 pGEM-T 载体。利用大肠杆菌 DH5 $\alpha$ ; 以及 Promega 出品的工具酶进行酶切、转化和序列测定。其中上述引物由上海博亚生物技术有限公司合成, 测序由日本 TAKARA 公司完成。所得结果见后附序列表中 SEQ ID NO: 1。

## 实施例 2

### 鸡 COL2A1 基因组 DNA 的提取

鸡 COL2A1 基因组 DNA 的提取采用 Wizard genomic DNA purification kit (Promega) 进行。具体操作见试剂盒说明书。提取 DNA 纯度测定 260/280 为 1.6-1.8 (BECKMAN, DU®640)。

### 鸡 COL2A1 基因组 DNA 的克隆

以 PgF 5' -GCT CGG AAG CAA GGC CCA CG -3'; PgR 5' -GCA CAA CGG CTC GGG CAA TGT GCT AAC G -3' 为上、下游引物, Takara LA Taq 酶 (GC buffer) 对全血中提取的 DNA 进

行 PCR 扩增。0.8% 琼脂糖电泳检查 PCR 结果，并采用凝胶回收试剂盒（GIBCO BRL）回收产物并连 pGEM-T 载体，酶切鉴定阳性克隆测序。

采用特异性引物对鸡 COL2A1 基因组 DNA 进行扩增，扩增产物特异性极高，将其插入 pGEM-T 载体，从测序结果可看出鸡 COL2A1 3' 端基因组 DNA 包含部分三螺旋区、3' 端肽及 3' 原肽，此克隆基因组部分含有内含子 19 个，外显子 20 个。内含子、外显子结构如表 2 示，所得鸡 COL2A1 3' 端基因的结构如图 2 所示。

表 2: 鸡 COL2A1 基因组 3' 端的外显子、内含子结构

外显子	编码区氨基酸	外显子 (bp)	内含子 (bp)
1		147	
2		243	112
3		188	594
4		289	407
5	964-999	108	79
6	946-963	54	134
7	910-945	108	109
8	892-909	54	268
9	856-891	108	97
10	838-855	54	160
11	802-837	108	309
12	766-801	108	92
13	712-765	162	268
14	694-711	54	113
15	676-693	54	108
16	640-675	108	89
17	622-639	54	94
18	604-621	54	82

注: 外显子、内含子的命名采用 Upholt 等 J Biol Chem, 1984, 259 (12): 7826-7834 的方法。

### 实施例 3

#### 鸡 COL2A1 $\alpha_1$ (II) 组织特异性表达

##### 鸡胚 COL2A1 表达谱

COL2A1 3' UTR 经 Blast 的核苷酸序列数据库检索为 COL2A1 所特有。因此以 COL2A1 cDNA 3' UTR 为扩增对象, 研究  $\alpha_1$  (II) 在发育的鸡胚中的表达。应用 RT-PCR 对 17 日龄鸡胚心、肝、玻璃体、角膜、皮肤、胸肌、胸骨等脏器进行分析, 所用引物为表 1 中的 col2a1-1F, col2a1-1R, 产物连 pGEMT-easy 载体, 酶切后测序。同时以 GAPDH 作为内参, 利用引物 PF<sub>GAPDH</sub> 5' GC AGA GGT GCT GCC CAG AAC 3'; PR<sub>GAPDH</sub> 5' TCA CTC CTT GGA TGC CAT GTG 3' 扩增 412bp 片段的 GAPDH。

从 17 日龄鸡胚 COL2A1 3' UTR RT-PCR 结果可以看出, 鸡 COL2A1 mRNA 在鸡胚的心、肝、玻璃体、角膜、皮肤、胸肌、胸骨、小肠、关节软骨、半月板、颅骨中有表达, 而脾脏、胸腺、睾丸却检测不到 COL2A1 mRNA。(参见图 3)。

##### 成鸡 II 型胶原蛋白表达谱

采用天然 II 型胶原检测试剂盒 (Chondrex) 对四周龄鸡心、肝、脾、肾、玻璃体、角膜、皮肤、胸肌、半月板、胰脏、胸腺、小肠、胃、睾丸、骨骼肌、大脑、小脑、关节软骨、胸骨、肺脏共 20 个脏器进行 ELISA 检测 II 型胶原蛋白的存在。(冻干组织各 5mg, 充分匀浆, 0.8ml 50mM 醋酸-0.2M NaCl (pH2.9-3.0) 溶解, 100 $\mu$ l 20mg/ml 的胃蛋白酶 (Sigma) 4 $^{\circ}$ C 消化 48h; 加入 250 $\mu$ l 10 $\times$ TSB, 并用 1 M NaOH

调 pH 至 8.0, 加入 100 $\mu$ l 2mg/ml 的弹性蛋白酶 (Sigma) 混匀, 4 $^{\circ}$ C 消化过夜。10000g, 4 $^{\circ}$ C 离心 5 分钟, 弃沉淀, 上清定容至 1ml, 4 $^{\circ}$ C 保存备用。

采用双抗体夹心 ELISA 法检测提取的 II 型胶原蛋白。选用两种针对鸡 II 型胶原蛋白分子不同表位的单克隆抗体, 其中, 一种作为包被抗体, 另一种作为酶标检测抗体。首先捕获抗体包被 ELISA 板, 4 $^{\circ}$ C 过夜, 洗涤三次, 处理好的样品 1: 200 稀释并与标准品一同点样于包被好的 ELISA 板, 37 $^{\circ}$ C 2h 洗涤同上; 生物素标记的鸡 II 型胶原蛋白单克隆抗体加入各孔, 37 $^{\circ}$ C 孵育 2h, 洗涤同上, 最后加入酶标链亲和素, 37 $^{\circ}$ C 1h; 充分洗涤, 最后加显色液 OPD, 37 $^{\circ}$ C 显色半小时, 1.25M 硫酸终止反应。测 OD490 光密度值。具体操作见试剂盒说明书。

成鸡 20 种脏器天然 II 型胶原蛋白检测实验发现, 鸡 II 型胶原蛋白在胸骨、关节软骨、胰脏、小肠有表达, 而在其它脏器却不见表达。(参见图 4)

同时以鸡天然 II 型胶原蛋白作为标准品, 制作浓度、吸光度标准曲线, 依据标准曲线计算相应各阳性组织中鸡 II 型胶原蛋白的表达量(参见图 5)。

据标准曲线得出胸骨中鸡 II 型胶原蛋白的表达量为 0.4%, 关节软骨表达量在四周鸡所检测组织中表达量最高达 1.5%, 胰脏、小肠中的表达量分别为 0.84% 与 0.2%。

#### 实施例 4 鸡 COL2A1 5' 原肽外显子 2 剪接分析

采用 RT-PCR 分析鸡 COL2A1 外显子 2 在鸡胚各脏器中的剪接情况, 以表 1 中 col2a1-5F, col2a1-5R 为上、下游引物对鸡皮肤、肝脏、眼球玻璃体、角膜、胸肌、小肠、胸骨、关节软骨等进行 PCR 扩增分析, 1% 琼脂糖电泳检测 PCR 扩增产物的特异性。RT-PCR 结果表明, 鸡 COL2A1 5' N 原肽外显子 2 存在于鸡心脏、肝脏、眼球玻璃体、角膜、胸肌、小肠, 而软骨组织如胸骨、关节软骨中以不含外显子 2 的表达形式存在。(参见图 6)



### 实施例 5 鸡 COL2A1 基因染色体的定位

为了将鸡 COL2A1 基因精细定位于染色体,采用染色体 ISH 技术,应用地高辛高效标记检测试剂盒 (Roche) 标记的鸡 COL2A1 3'UTR 基因片段为探针,与中期分裂相的鸡染色体进行杂交,研究其在染色体上的定位。

采用 ISH 原位杂交,以地高辛标记的鸡 COL2A1 3'UTR 基因片段为探针,与鸡中期分裂相的染色体精细杂交,杂交信号经统计分析,将鸡 COL2A1 基因定位于 4 号染色体短臂 2 区 (参见图 7, 8, 和 9)。

### 实施例 6 鸡 COL2A1 同源性比较

利用 Dnastar 软件包中的 MegAlign 对已获得的鸡 COL2A1 与从 GenBank 数据库中检索到的犬 (AF023169, AF242201)、人 (MM001844)、斑点鱼 (U23822)、小鼠 (M65161) 等不同种属的 II 型胶原三螺旋区 cDNA 序列及蛋白质序列进行同源性的比较和分析, Genedoc 程序编辑, 绘制进化树 (参见图 10)。

鸡 COL2A1 cDNA 序列与已知的人、犬、小鼠、斑点鱼进行了同源性比较, 结果显示鸡 COL2A1 与犬 CII 在种族进化上同源性最高, 其氨基酸与 cDNA 同源性分别为 79.03 %, 94.77 %; 其次为鸡与人, 同源性分别为 78.96 %, 93.89 %; 鸡与鼠同源性分别为 77.38 %, 92.90 %。人与犬在所比较的五种种属中同源性最高, 氨基酸与蛋白质分别为 91.89 %、98.52 %。其次为人与鼠, 分别为 88.63 %, 96.15 %, 比较结果建立进化树。

## 参考文献

1. Jenkins JK, Hardy KJ. Biological modifier therapy for the treatment of rheumatoid arthritis. *Am J Med Sci*, 2002,323(4):197-205.
2. Danos O, Malligan MC. Safe and efficient generation of recombinant retrovirus with amphotropic host ranges. *Proc Natl Acad Sci USA*, 1988,85:6460-6465.
3. Roessler BJ, Allen ED, Wilson JM. Adenoviral-mediated gene transfer to rabbit synovium in vivo. *J Clin Invest*, 1993,92:1085-1092.
4. Trentham DE, Dynesius-Trentham RA, Orav EJ, et al. Effects of oral administration of type II collagen on rheumatoid arthritis. *Science*, 1993,261: 1727-1730,
5. Sandell LJ, Prentice HL, Kravis D, Upholt WB. Structure and sequence of the chicken type II procollagen gene . *J Biol Chem* , 1984,259 (12) 7826-7834.
6. Horton RM, Hunt HD, Ho SN, et al. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene*, 1989, 77:61-68.
7. Sambrook J, Fritsch EF, Maniatis T. *Molecular cloning: a laboratory manual*. 2<sup>nd</sup> ed Cold Spring Harbor Laboratory Press, 1989.
8. Nah DH, Upholt WB. Type II collagen mRNA containing an alternatively spliced exon predominates in the chick limb prior to chondrogenesis. *J Biol Chem*, 1991, 266 34:23446-23452.
9. Rousseau JC, Farjanel J, Boutillon MM, et al. Processing of type XI collagen. Determination of the matrix forms of the alpha 1 (XI) chain. *J Biol Chem*, 1996,271(39): 23743-8.
10. Snellman A, Keranen MR, Hagg PO, et al. Type XIII collagen forms homotrimers with three triple helical collagenous

domains and its association into disulfide-bonded trimers is enhanced by proly 4-hydroxylase. *J Biol Chem*, 2000, 275(12):8936-44.

11. Young MF, Vogeli G, Nunez AM, et al. Isolation of cDNA and genomic DNA clones encoding type II collagen. *Nucleic Acids Res*, 1984, 12 (10): 4207-4228.

12. Marshall GE, Konstas AGP, Lee WR. Collagens in ocular tissues. *BrJ Ophthalmol*, 1993, 77:515-524.

13. Seery CM, Davison PF. Collagen of the bovine vitreous. *Invest Ophthalmol Vis Sci*, 1991, 32:1540-1550.

14. Huerre-Jeanpierre C, Mattei MG, Weil D, et al. Further evidence for the dispersion of the human fibrillar collagen genes. *Am J Hum Genet*, 1986, 38(1): 26-37.

15. Ausar SF, Beltramo DM, Castagna LF, et al. Treatment of rheumatoid arthritis by oral tolerance of bovine tracheal type II collagen. *Rheumatol Int*, 2001, 20 :138-144.

16. Barnett ML, Combitchi D, Trentham DE. A pilot trial of oral type II collagen in the treatment of juvenile rheumatoid arthritis. *Arthritis & Rheumatism*, 1996, 39 4:623-628.

17. Kim WU, Lee WK, Ryoo JW, et al. Suppression of collagen-induced arthritis by single administration of poly(lactic-co-glycolic acid) nanoparticles entrapping type II collagen: a novel treatment strategy for induction of oral tolerance. *Arthritis Rheum*, 2002, 46:1109-20.

# SEQUENCE LISTING

<110> 中国人民解放军军事医学科学院附属医院

<120> 一种编码鸡 II 型胶原蛋白的全长多核苷酸序列及其用途

<130> IDC020013

<160> 3

<170> PatentIn version 3.1

<210> 1

<211> 5495

<212> DNA

<213> genomic DNA

<400> 1

```

ccaggcaagg atggcgacag tgtaagtggg gcacggccat ggggtgggct ggcaaaggat   60
ggtcacagag accacatcct catctcctc tctctcccat aggggtctgac ggggtccatt   120
ggccccctg gccctgctgg cccaacggg gagaagggtga gagcagcatc acagcacccc   180
acatlacgcc ccatgggatg accccagtgc ctccacctct ccatcctttc ttttcagggg   240
lgaatccggc cctcctgggc catctgggtg tgccgggtgcc cgtgggtgcc ccgtaagcac   300
aatgtctgca gcccttgggt gccctaacc ttcacctaa acccccatca acccctttat   360
caacctcccc catctcttcc cattaggggt agcgtggcga gcccggtgcc cccggtcctg   420
ctggatttgc tggcccccg gtgagtgtt caccgccgaag ccccatcgc acaccacgt   480
cttcaccca catctcacc cactcatgg tggtgtgtgt tccatcagg gcgccgatgg   540
acaacccggg gccaaaggcg agcagggaga gcccgggcag aagggtgacg cgggcgctcc   600
tggtcccaa ggctcctccg gcgtcctgg ccccaggta caacacaaa tggggcaaac   660
cccaaattt gggacgtcac gggcccaatg caggcacact gcagctccg ttcggatttg   720
taacctgtt ttctctctt ctaggggtcc aaccgggtgc actggcccc aaggagctcg   780
tggggctcag ggccccctg tgagtaccgg ggggtgggct gcagggtggg gaaggagcgg   840

```

ccgtggggct gagctgtgtc tgagccgttt ctctctccc tctctctct gactctgtga 900  
 ttccctcccc agggagccac gggattcccc ggagctgccg gccgtgtggg accgcccggc 960  
 cctaattgtga glctgggggc gttctgggat tgccccacc tggggtttgg gcgctgcttc 1020  
 cccgcgtgc gtgttgagg gggcactgtt tcctgcaca gacacgtggg gttttctcc 1080  
 ttggctctct gatgttggt tttggggcca ttccaatggt agagaaggac ttttctaagg 1140  
 gcaagagctc cccaagaagc agcagtggga tgcgggtgat aaagatggaa tggctgcctc 1200  
 tggtttgac caacgtgct ttccttcct ttagggtaac ccaggcccc ccggaccccc 1260  
 cggtctgct ggcaaagacg gcccgaagg tggtcgtggc gacgccggcc ccccggccg 1320  
 tgcaggtgac cccggcctcc aaggccccgc cggccccccc ggcgagaagg gcgaaccogg 1380  
 cgaggacggc cccgcgtga ggattctggg ggtctcctcc ctccgtgcac cccctggctg 1440  
 cgtggtgccg ttgltcttag tctgatttcc ccctctgtg cccgcaggg tcccacggc 1500  
 cccccccggc cctcaaggct tggcaggaca gcgtggtatt gtgggtctcc caggacagcg 1560  
 tggtagaga ggcttccccg gactgccggg gccatcggtg aglgggtcgc lctcatttgg 1620  
 gtgcactgaa tcctatgggg tgcagagatg tgggggccgc gatgctctgg agcccatctc 1680  
 aggggtgcc agcccttttg tgcagccgg ggacaccgtt tgcaggtggg ttggggtttt 1740  
 gcggagctcc ttttcccca ccaggagccg ctggtgcaag gcttaaagcc ggggcaggaa 1800  
 aaccatcagl ggltatttgt tgcagagggg tctgggagcc ataaaaaac gggaaggggc 1860  
 agcgtgggg tctctccac tcatgcacct cttcccatc tttcaggag aacctggaaa 1920  
 gcaaggagcg cctggctctg cgggtgaccg aggtcccccc ggccccgtgg gccccctgg 1980  
 gctgacaggl cctgctggag aaccggggcg cgaggaagc aaaacccac agcatcacag 2040  
 cggcaccggg catcaccaac cccatggcac agctcagctc ccagagctcc ccggtgtctt 2100  
 tttctccagc actgaaagga gactttgcac aaatcctgct ccaccgggt tgtaacatcc 2160

ccttttctc ctagggcaac cctggtgctg acggtcccc aggcagggat ggcgagctg 2220  
 gcgtgaaggt gagcttgcca tgcgtcccc attggcactc gccatccccg tgccaaaagc 2280  
 tglggggttt tgcacagatc tgacctctct gttgtctgct cgcagggtga tcgtggtgag 2340  
 accggccctg tgggtgctcc cgggtctcct ggagcccctg gcgcccccg ccctgttggt 2400  
 cccacigga aacaaggaga cagaggcgag acggtgagtg ctggcacaag ggtttagggt 2460  
 ttagggtctc cttatggctg aaaatgigca ggggttcccc tcaaggttg ttccttgac 2520  
 cagtgtgag tgcattttaa gatgctgtga ggcaccaaca gctgctgatt gtcactgttg 2580  
 cccggatctg ggggtcggag catggggctg gctcagacac ccccgaaatc ccaaattcat 2640  
 ggcttcgagg tgggtcttct ggtcgtggc accttctgat gtcctttttt tctccctgca 2700  
 ggggtcacaa gggcccatgg gtccctctgg tcccgtgga gctcaggaa tgccggtgag 2760  
 tgggtgctgag tgcctcggca catccacgt acagagcgtg gggctcctgcg tgccaggagg 2820  
 gggcttgcca cccigagccc gacacagccc tgtcccact ttagggctcc caaggacctc 2880  
 gtggtgacaa aggtgagacg ggagaggctg gagagagagg gctgaagggc caccgcggct 2940  
 tcaccggtct gcagggtctg cccggaccac ccgtaagttg gtttggggag cactgagccc 3000  
 cccccccgt acgatcggc tcctttgggg tctctgtggc caccgaggct ctgtctggcc 3060  
 caaagtgtg accgcagagc tgtgaccacc ccggcttct cctcagggcc cgtctggaga 3120  
 ccaagggtgct gccggtcctg ctggctcctc cggctccaga gtaagtctg acggtggtgt 3180  
 ttgggggtgt ggaaggggaa ggagcagcag tggcctcctt gggcacctgc agcctctgtt 3240  
 cgctcctgtc tgctcatcag caccatcgcc ttccctgccc tgaggccccg caatgccttc 3300  
 acctccccgt ttggggctc tctctagggt tccccctggt cccgtcggcc cctctggcaa 3360  
 agacggctcl aacggcatgc ccggcccat cggctcctcc ggtccccgtg gacggagtgg 3420  
 tgaaccggc cctgcggtga gtcctggtga ggggaggcag ggaatggggt ccagctcgca 3480  
 gagcagccca tcagcatcac ttctttctcc catagggtcc lcctggaaac cccggtcctc 3540

ccggtccicc tggcccccc ggcaccggca tcgacatgtc tgcttttgct ggactgggtc 3600  
agacggagaa gggccccgac cccatccgct acatgggggc agacgaggcg gccggagggc 3660  
tgcggcagca cgacgtggag gtggacgcca ccctcaaatc cctcaacaat cagattgaga 3720  
gcatccgcag ccccgagggc lccaagaaga accctgccag gacctgccgc gacatcaaac 3780  
tctgccatcc cgagtggag agcggtaaga gctccgcgtg cctctcccgt cctcccctct 3840  
tccccacagg agagcatccc cagcgtcctc gcaccgacct gcggtcaggt tggatgttag 3900  
gaaagaitcc ttgtccaaaa gagctctggg cgctgggctg ggctgcccgg ggaggtgggg 3960  
caglcgtgl ccccataggt gttggggaac tgtggagatg tggcacttgg gagcgtggct 4020  
tagtggggat gaggcagcag ttggaccaat cttcgaggtc ttclccagtc llaatggctc 4080  
tgtgttctg tcggtgtgca tgggtgtgat ggglggccat ttagacttgg cgatctttga 4140  
ggtcttllcc gatcttaacg actcctagac clccccaacc ccatgaacgc tgtttgtcct 4200  
ccccctgca ggagattact ggattgacct gaaccagggc tgcaccttgg acgcatcaa 4260  
agtattctgc aacatggaga caggcgagac ctgcgltac ccgaccccca gcagcatccc 4320  
caggaagaac tgggtggacca gcaagacgaa agacaagaag cacgtctggt ttgcagagac 4380  
catcaacggc ggtttccacg tgggtgtccc ccgggtgtcc ttggaaggat cgatcccacc 4440  
tgggalgtcc ticttgcggt catgtggatg ggttttaatg aagtatataga gggatgattct 4500  
gaaggtgtag gtttgggtca gttcagctcc acaaatcaaa gggaaaggat gggatggagc 4560  
aactgagctc cctcggtttg tttggcccag aaaagggtgag gatgagggga ggcctcacgg 4620  
ccttacagcc cctlacggcc ctacagcagc gttaggaaaa aagtictgcc ccggagctgt 4680  
gttgggcaca gaacagccct gtgatccgg agctcgggga gcattgggac aacgctctca 4740  
gacatlggtt ttgggtcagg tcttggttaa cgtgatgtgc agggggcaac cagcccatgg 4800  
gtgggcttta aggacccttc caagccaacc attccatggt tctgtgatct gtaaggacct 4860

ttccaatcca aaccactctg attttttct cagccatttg ggaacctgaa gtacggaagt 4920  
cctcccaaaa agctcctgag agtaagggtg tcataatgcc cgcaggctll aactcctcac 4980  
ctcttccctc cagttcagct acggcgatga gaacctgtcc cccaacaccg ccagcatcca 5040  
gatgaccttc ctgcgcctcc tgtccaccga gggctcccag aacgtcacct accactgcaa 5100  
gaacagcatc gcctacatgg acgaggagac gggcaacctg aagaaagcca tcctcatcca 5160  
gggatccaac gacgtggaga tcagagccga gggcaacagc aggttcacct acagcgtctt 5220  
ggaggacggc tgcacggtag gttgctgggc gcctgcaaag gaaaggtgca gatggggagg 5280  
gggaggctga ggctgggggg atgaggccgg agcagclgac agcatccctg ccctccttcc 5340  
ctccccagaa acacactggc aaalggggca agacgggtgat cgagtaccgg tcgcagaaga 5400  
cctcgcgcct gccaltgta gatattgcac ctatggacat tggcggagcc gatcaggagt 5460  
ttggcgtgga lattggccca gtctgcttct lgtaa 5495

<210> 2  
<211> 4793  
<212> DNA  
<213> cDNA

<400> 2  
atgcacggcc gccgcccgcc ccgctccgcc gctctctcc tctctctct ccttcicacg 60  
gccgcccga cgcgcagga ccgcgacct cgacaacctg gcccgaagg acagaaggga 120  
gaaccggag atattaaaga tgtttagga ccccgagggc ctccaggacc acagggccca 180  
gcaggagagc agggacagcg aggggacctt ggcgagaagg gggagaagg tgctcctggc 240  
ccccgtggga gggatggaga acccggcacc cctggaaacc caggcccccc cgttcccccc 300  
ggacctcctg gcccccccgg acilgggtgga aactttgcgg cgcagatggc gggcggcttc 360  
galgagaagg cgggtggagc gcagatgggt gtcatgcagg gacctatggg ccctatggga 420  
ccccgcggcc cccctggccc cactggcgca cctgggtccc agggatttca aggcaacccc 480



ggtgagcccg gcgaaccccg cgctgctggt ccgatgggtc cccggggacc tccgggacca 540  
 cctgggaaac ccggtgacga tggtagaca ggcaaaccg gcaaatctgg tgaacglggc 600  
 cccccggcc cccagggcgc tcgtggclic cctgggactc ctggtctccc cggagtgaag 660  
 ggccaccgag gctaccccg tttggatggt gccaaaggag aggcgggggc tcctggagcc 720  
 aagggtgaat cgggttcacc gggtagaac ggctccccg gcccctggg acccgtggg 780  
 ctgcccggag agcaggagc tcccggcccc tccggcgccg ccggtgctcg tggcaatgac 840  
 ggtctcccig gccctgctgg accccctgga cccgtcggcc ctgcccggag ccccggttc 900  
 cccggagccc ccggtlcaaa gggtagagc ggccccactg gtgcacgggg tcccgagggt 960  
 gcccaaggac cccgcggcga atccggcacc cccggtctc ccggccccgc tggcgaccc 1020  
 ggtaaccag ggactgatgg catccccggg gccaaagggt cggcggggtc cccgggcatt 1080  
 gcaggcgctc caggattccc cggcccacgc ggcccccccg gacccaagg tgccaccgga 1140  
 cacttgggac ccaaaggaca gacgggcgaa cccggcatcg caggcttcaa gggcgagcaa 1200  
 ggaccgaagg gcgagacggg cccgcagga cccaagggtg ccccgggcc ggctggtgag 1260  
 gaaggcaaga gaggagctcg tggtagacct ggtgccgccc gccctgtggg ccccccgga 1320  
 gaaagggcg clcctggcaa ccgtggattc cccgggcagg acgggctggc cggaccaag 1380  
 ggtgctccag gtgaacgcgg ccccgctggt ctgcgggtc ccaaagggtc caccggtgac 1440  
 cccggacgtc ccggagagcc cgggctgccc ggagcgaggg gtctcaccgg ccgccccgc 1500  
 gatgcccggac ctcaaggcaa agtcggccca actggtgctc ctggcgagga tggccgcccc 1560  
 ggcccccccg gacctcaggg tgctcgtggg cagcctggtg tgatgggttt ccccggtccc 1620  
 aaaggcgcta atggtagacc tggaaaagct ggagagaaag gactgcccgg cggcccaggg 1680  
 ctcgggggtc tgcctggcaa ggatggggag acgggagctg ccggcccccc tggaccgct 1740  
 ggtcctglgg glgagagagg agagcaagga gccccggtc cttccggtt ccagggactg 1800  
 cccggnccc cagggtcccc tggggagagc ggcaaaccg gagaccaggg tgttcctgga 1860

gaagccggtg cccccgtct tgttggtccc agaggtgaac gtggattccc cggatgaacgc 1920

ggctctcccg gtgccaagg gctgcagggt ccccggtggc tccccggaac gcccggcact 1980

gacggaccca aggggtgcaac cgggtccagcc ggccccaacg gtgcccagggt tccccagggt 2040

ctgcaggga tggccggtga gagaggagca gctggcatcg ctggcctcaa ggggtgaccgg 2100

ggagatgttg gtgagaaagg acctgaggga gctccaggca aggatggcgc acgtggtctg 2160

acgggtccca ttggtcccc tggccctgct ggccccaacg gtgagaagggt tgaatccggc 2220

cctcctggtc catctggtgc tgccggtgcc cgtggtgccc ccggtgagcg tggcgagccc 2280

ggtgcccccg gtctgtctgg atttctggc cccccggcg ccgatggaca acccggtgcc 2340

aaaggcgagc agggagagcc cgggcagaag ggtgacgcgg gcgctcctgg tccccagggt 2400

ccctccggcg ctcttgcccc ccagggccca accggtgtca ctggtccaa aggagctcgt 2460

ggggtcagg gtccccctgg agccacggga ttccccggag ctgccggccg tgtgggaccg 2520

cccgcccta atgtaacct agggcccccc ggacccctg gctctgtgg caaggacggc 2580

ccaagggtg ttctgtggca cgccggcccc cccggccglt caggtgacct cggcctccaa 2640

ggccccccg gcccccccg cgagaagggt gaacccggcg aggacggccc cgcgggtccc 2700

gacggcccc ccggccctca aggcctggca ggacagcgtg gtattgtggg tctcccagga 2760

cagcgtgglt agagaggctt ccccgactg ccggggccat cgggagaacc tggaaagcaa 2820

ggagcgctg gctctgcggg tgaccgaggt cccccggcc ccgtggggcc ccctgggctg 2880

acgggtcctg ctggagaacc cgggcgcgag ggcaaccctg gtgctgacgg tctcccaggc 2940

agggatggcg cagctggcgt gaagggtgat cgtggtgaga cgggccctgt ggggtgcccc 3000

ggtgclcctg gagccccctg cggccccggc cctgttggtc cactggaaa acaaggagac 3060

agaggcgaga cgggtgcaca agggcccatg ggtccctctg gtcccgctgg agctcgagga 3120

atgccggglt cccaaggacc tcgtggtgac aaaggtgaga cgggagaggc tggagagaga 3180

gggctgaagg gccaccgtgg cttcaccggt ctgcagggtc tgcccggacc acccggcccg 3240  
 tctggagacc aagggtgctgc cgggtcccgct ggtccclccg gtcccagagg tccccctggt 3300  
 cccgtcggcc cctctggcaa agatggclct aacggcatgc ccggcccat cggtcctccc 3360  
 ggtccccgtg gacggaglgg tgaaccggc cctgcgggtc ctcttgaaa ccccggtcct 3420  
 cccggtcctc ctggccccc cggcacggc atcgacatgt ctgcttttgc tggactgggt 3480  
 cagacggaga agggccccga ccccatccgc tacatgaggg cagacgaggc ggccggaggg 3540  
 ctgcggcagc acgacgtgga ggtggatgcc accctcaa at ccctcaaca tcagattgag 3600  
 agcatccgca gccccgagg ctccaagaag aacctgcca ggacctgcc cgacatcaaa 3660  
 ctctgccatc ccgaglggaa gagcggagat tactggattg acccgaacca gggctgcacc 3720  
 ttggacgcca tcaaagtatt ctgcaacatg gagacggcg agacctgcgt ctacccgacc 3780  
 cccagcagca tccccaggaa gaactggtgg accagcaaga cgaaagaca gaagcacgtc 3840  
 tggtttgag agaccatcaa cggcggtttc cacttcagct acggcgatga gaacctgtcc 3900  
 cccaacaccg ccagcatcca gatgaccttc ctgcgcctcc tgtccaccga gggctcccag 3960  
 aacgtcacct accactgcaa gaacagcatc gcctacatgg acgaggagac gggcaacctg 4020  
 aagaaagcca lcctcatcca gggatccaac gacgtggaga tcagagccga gggcaacagc 4080  
 aggtlcacct acagcgtctt ggaggacggc tgcacgaaac aactggcaa atggggcaag 4140  
 acggtgatcg agtaccggtt gcagaagacc tcgcgcctgt ccattgtaga tactgcacct 4200  
 atggacattg gcggagccga tcaggagttt ggcgtggata ttggcccagt ctgcttcttg 4260  
 taaaaagggg tgttggtatt tgtgtgttg ttgtgtgtt ggttggtgtt tttgtttct 4320  
 ttttttttt tttttagaaa agaaaggaat ccagcccaat ccataaaag caaaccagtc 4380  
 ccacccccag gaccgcacg ttcccagcac aacttctgca ctgaacggat ggcacgaccc 4440  
 cgcgccctt cgggacccic cggcgcgtc accgggcaga ctgcgaaata caaccacggg 4500  
 cttatattta ttlatcgct tcctggaagg cctgglttcg lagggcgggt ggaggtggga 4560

atcaatctgg caggtgtgac ggccccctc cccacaaagg gatctggcaa acgcaggtat 4620  
 cgcgaaatccc ctccccctccc cgtgtatcac cagcaggagt gctaatgtat catacaacag 4680  
 aaatgggtgcl attcttgtaa aacaagtctg tattttttaaa catcagttga tataaaaaca 4740  
 acaaaaaaaaa aaacttttgg tggaaagtaa aaaaaacaaa aaaaaaaaaaaa aaa 4793

<210> 3  
 <211> 1420  
 <212> PRT  
 <213> chicken

<400> 3

Met His Gly Arg Arg Pro Pro Arg Ser Ala Ala Leu Leu Leu Leu Leu  
 1 5 10 15

Leu Leu Leu Thr Ala Ala Ala Ala Ala Gln Asp Arg Asp Leu Arg Gln  
 20 25 30

Pro Gly Pro Lys Gly Gln Lys Gly Glu Pro Gly Asp Ile Lys Asp Val  
 35 40 45

Val Gly Pro Arg Gly Pro Pro Gly Pro Gln Gly Pro Ala Gly Glu Gln  
 50 55 60

Gly.Gln Arg Gly Asp Arg Gly Glu Lys Gly Glu Lys Gly Ala Pro Gly  
 65 70 75 80

Pro Arg Gly Arg Asp Gly Glu Pro Gly Thr Pro Gly Asn Pro Gly Pro  
 85 90 95

Pro Gly Pro Pro Gly Pro Pro Gly Pro Pro Gly Leu Gly Gly Asn Phe  
 100 105 110

Ala Ala Gln Met Ala Gly Gly Phe Asp Glu Lys Ala Gly Gly Ala Gln  
 115 120 125

Met Gly Val Met Gln Gly Pro Met Gly Pro Met Gly Pro Arg Gly Pro  
 130 135 140

Pro Gly Pro Thr Gly Ala Pro Gly Pro Gln Gly Phe Gln Gly Asn Pro  
 145 150 155 160

Gly Glu Pro Gly Glu Pro Gly Ala Ala Gly Pro Met Gly Pro Arg Gly  
 165 170 175

Pro Pro Gly Pro Pro Gly Lys Pro Gly Asp Asp Gly Glu Thr Gly Lys  
 180 185 190

Pro Gly Lys Ser Gly Glu Arg Gly Pro Pro Gly Pro Gln Gly Ala Arg  
 195 200 205

Gly Phe Pro Gly Thr Pro Gly Leu Pro Gly Val Lys Gly His Arg Gly  
 210 215 220

Tyr Pro Gly Leu Asp Gly Ala Lys Gly Glu Ala Gly Ala Pro Gly Ala  
 225 230 235 240

Lys Gly Glu Ser Gly Ser Pro Gly Glu Asn Gly Ser Pro Gly Pro Met  
 245 250 255

Gly Pro Arg Gly Leu Pro Gly Glu Arg Gly Arg Pro Gly Pro Ser Gly  
 260 265 270

Ala Ala Gly Ala Arg Gly Asn Asp Gly Leu Pro Gly Pro Ala Gly Pro  
 275 280 285

Pro Gly Pro Val Gly Pro Ala Gly Ala Pro Gly Phe Pro Gly Ala Pro  
 35

290

295

300

Gly Ser Lys Gly Glu Ala Gly Pro Thr Gly Ala Arg Gly Pro Glu Gly  
305 310 315 320

Ala Gln Gly Pro Arg Gly Glu Ser Gly Thr Pro Gly Ser Pro Gly Pro  
325 330 335

Ala Gly Ala Pro Gly Asn Pro Gly Thr Asp Gly Ile Pro Gly Ala Lys  
340 345 350

Gly Ser Ala Gly Ala Pro Gly Ile Ala Gly Ala Pro Gly Phe Pro Gly  
355 360 365

Pro Arg Gly Pro Pro Gly Pro Gln Gly Ala Thr Gly Pro Leu Gly Pro  
370 375 380

Lys Gly Gln Thr Gly Glu Pro Gly Ile Ala Gly Phe Lys Gly Glu Gln  
385 390 395 400

Gly Pro Lys Gly Glu Thr Gly Pro Ala Gly Pro Gln Gly Ala Pro Gly  
405 410 415

Pro Ala Gly Glu Glu Gly Lys Arg Gly Ala Arg Gly Glu Pro Gly Ala  
420 425 430

Ala Gly Pro Val Gly Pro Pro Gly Glu Arg Gly Ala Pro Gly Asn Arg  
435 440 445

Gly Phe Pro Gly Gln Asp Gly Leu Ala Gly Pro Lys Gly Ala Pro Gly  
450 455 460

Glu Arg Gly Pro Ala Gly Leu Ala Gly Pro Lys Gly Ala Thr Gly Asp  
465 470 475 480

Pro Gly Arg Pro Gly Glu Pro Gly Leu Pro Gly Ala Arg Gly Leu Thr  
485 490 495

Gly Arg Pro Gly Asp Ala Gly Pro Gln Gly Lys Val Gly Pro Thr Gly  
500 505 510

Ala Pro Gly Glu Asp Gly Arg Pro Gly Pro Pro Gly Pro Gln Gly Ala  
515 520 525

Arg Gly Gln Pro Gly Val Met Gly Phe Pro Gly Pro Lys Gly Ala Asn  
530 535 540

Gly Glu Pro Gly Lys Ala Gly Glu Lys Gly Leu Pro Gly Ala Pro Gly  
545 550 555 560

Leu Arg Gly Leu Pro Gly Lys Asp Gly Glu Thr Gly Ala Ala Gly Pro  
565 570 575

Pro Gly Pro Ala Gly Pro Val Gly Glu Arg Gly Glu Gln Gly Ala Pro  
580 585 590

Gly Pro Ser Gly Phe Gln Gly Leu Pro Gly Pro Pro Gly Pro Pro Gly  
595 600 605

Glu Ser Gly Lys Pro Gly Asp Gln Gly Val Pro Gly Glu Ala Gly Ala  
610 615 620

Pro Gly Leu Val Gly Pro Arg Gly Glu Arg Gly Phe Pro Gly Glu Arg  
625 630 635 640

Gly Ser Pro Gly Ala Gln Gly Leu Gln Gly Pro Arg Gly Leu Pro Gly  
645 650 655

Thr Pro Gly Thr Asp Gly Pro Lys Gly Ala Thr Gly Pro Ala Gly Pro  
660 665 670

Asn Gly Ala Gln Gly Pro Pro Gly Leu Gln Gly Met Pro Gly Glu Arg  
675 680 685

Gly Ala Ala Gly Ile Ala Gly Leu Lys Gly Asp Arg Gly Asp Val Gly  
690 695 700

Glu Lys Gly Pro Glu Gly Ala Pro Gly Lys Asp Gly Ala Arg Gly Leu  
705 710 715 720

Thr Gly Pro Ile Gly Pro Pro Gly Pro Ala Gly Pro Asn Gly Glu Lys  
725 730 735

Gly Glu Ser Gly Pro Pro Gly Pro Ser Gly Ala Ala Gly Ala Arg Gly  
740 745 750

Ala Pro Gly Glu Arg Gly Glu Pro Gly Ala Pro Gly Pro Ala Gly Phe  
755 760 765

Ala Gly Pro Pro Gly Ala Asp Gly Gln Pro Gly Ala Lys Gly Glu Gln  
770 775 780

Gly Glu Pro Gly Gln Lys Gly Asp Ala Gly Ala Pro Gly Pro Gln Gly  
785 790 795 800

Pro Ser Gly Ala Pro Gly Pro Gln Gly Pro Thr Gly Val Thr Gly Pro  
805 810 815

Lys Gly Ala Arg Gly Ala Gln Gly Pro Pro Gly Ala Thr Gly Phe Pro  
820 825 830



45

Gly Ala Ala Gly Arg Val Gly Pro Pro Gly Pro Asn Gly Asn Pro Gly  
835. 840 845

Pro Pro Gly Pro Pro Gly Ser Ala Gly Lys Asp Gly Pro Lys Gly Val  
850 855 860

Arg Gly Asp Ala Gly Pro Pro Gly Arg Ala Gly Asp Pro Gly Leu Gln  
865 870 875 880

Gly Pro Ala Gly Pro Pro Gly Glu Lys Gly Glu Pro Gly Glu Asp Gly  
885 890 895

Pro Ala Gly Pro Asp Gly Pro Pro Gly Pro Gln Gly Leu Ala Gly Gln  
900 905 910

Arg Gly Ile Val Gly Leu Pro Gly Gln Arg Gly Glu Arg Gly Phe Pro  
915 920 925

Gly Leu Pro Gly Pro Ser Gly Glu Pro Gly Lys Gln Gly Ala Pro Gly  
930 935 940

Ser Ala Gly Asp Arg Gly Pro Pro Gly Pro Val Gly Pro Pro Gly Leu  
945 950 955 960

Thr Gly Pro Ala Gly Glu Pro Gly Arg Glu Gly Asn Pro Gly Ala Asp  
965 970 975

Gly Leu Pro Gly Arg Asp Gly Ala Ala Gly Val Lys Gly Asp Arg Gly  
980 985 990

Glu Thr Gly Pro Val Gly Ala Pro Gly Ala Pro Gly Ala Pro Gly Ala  
995 1000 1005

Pro Gly Pro Val Gly Pro Thr Gly Lys Gln Gly Asp Arg Gly Glu

1010	1015	1020
Thr Gly Ala Gln Gly Pro Met Gly Pro Ser Gly Pro Ala Gly Ala		
1025	1030	1035
Arg Gly Met Pro Gly Pro Gln Gly Pro Arg Gly Asp Lys Gly Glu		
1040	1045	1050
Thr Gly Glu Ala Gly Glu Arg Gly Leu Lys Gly His Arg Gly Phe		
1055	1060	1065
Thr Gly Leu Gln Gly Leu Pro Gly Pro Pro Gly Pro Ser Gly Asp		
1070	1075	1080
Gln Gly Ala Ala Gly Pro Ala Gly Pro Ser Gly Pro Arg Gly Pro		
1085	1090	1095
Pro Gly Pro Val Gly Pro Ser Gly Lys Asp Gly Ser Asn Gly Met		
1100	1105	1110
Pro Gly Pro Ile Gly Pro Pro Gly Pro Arg Gly Arg Ser Gly Glu		
1115	1120	1125
Pro Gly Pro Ala Gly Pro Pro Gly Asn Pro Gly Pro Pro Gly Pro		
1130	1135	1140
Pro Gly Pro Pro Gly Thr Gly Ile Asp Met Ser Ala Phe Ala Gly		
1145	1150	1155
Leu Gly Gln Thr Glu Lys Gly Pro Asp Pro Ile Arg Tyr Met Arg		
1160	1165	1170
Ala Asp Glu Ala Ala Gly Gly Leu Arg Gln His Asp Val Glu Val		
1175	1180	1185

Asp Ala Thr Leu Lys Ser Leu Asn Asn Gln Ile Glu Ser Ile Arg  
1190 1195 1200

Ser Pro Glu Gly Ser Lys Lys Asn Pro Ala Arg Thr Cys Arg Asp  
1205 1210 1215

Ile Lys Leu Cys His Pro Glu Trp Lys Ser Gly Asp Tyr Trp Ile  
1220 1225 1230

Asp Pro Asn Gln Gly Cys Thr Leu Asp Ala Ile Lys Val Phe Cys  
1235 1240 1245

Asn Met Glu Thr Gly Glu Thr Cys Val Tyr Pro Thr Pro Ser Ser  
1250 1255 1260

Ile Pro Arg Lys Asn Trp Trp Thr Ser Lys Thr Lys Asp Lys Lys  
1265 1270 1275

His Val Trp Phe Ala Glu Thr Ile Asn Gly Gly Phe His Phe Ser  
1280 1285 1290

Tyr Gly Asp Glu Asn Leu Ser Pro Asn Thr Ala Ser Ile Gln Met  
1295 1300 1305

Thr Phe Leu Arg Leu Leu Ser Thr Glu Gly Ser Gln Asn Val Thr  
1310 1315 1320

Tyr His Cys Lys Asn Ser Ile Ala Tyr Met Asp Glu Glu Thr Gly  
1325 1330 1335

Asn Leu Lys Lys Ala Ile Leu Ile Gln Gly Ser Asn Asp Val Glu  
1340 1345 1350

Ile Arg Ala Glu Gly Asn Ser Arg Phe Thr Tyr Ser Val Leu Glu  
1355 1360 1365

Asp Gly Cys Thr Lys His Thr Gly Lys Trp Gly Lys Thr Val Ile  
1370 1375 1380

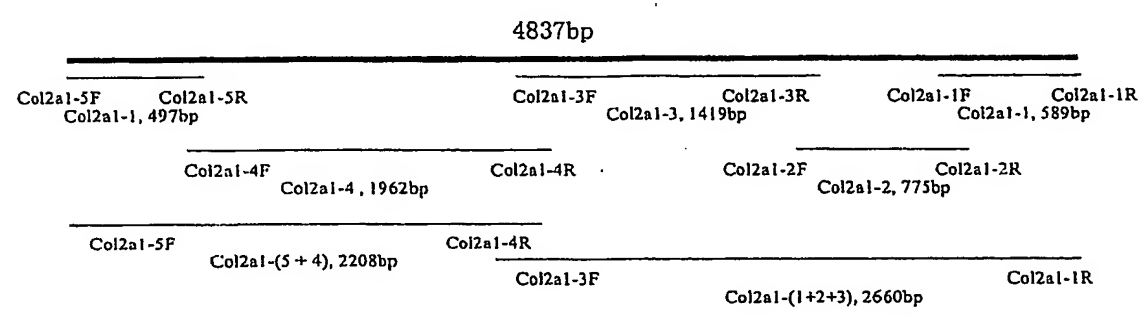
Glu Tyr Arg Leu Gln Lys Thr Ser Arg Leu Ser Ile Val Asp Thr  
1385 1390 1395

Ala Pro Met Asp Ile Gly Gly Ala Asp Gln Glu Phe Gly Val Asp  
1400 1405 1410

Ile Gly Pro Val Cys Phe Leu  
1415 1420

# 说明书附图

图 1



鸡 II 型胶原 PCR 克隆策略

图 2

信号肽

N-原肽

MHGRPPPSAAILLLILLITAAAAA

QDRDLRQPGPKQKGEPGDIKDVVGRGPPGPQQPAGEQQQRGRGEKGEKAGPGRGRDGEPTGPNPGPPGPPGPPGLGN

FAA QMAGGFDEKAGGAQMGMQ GPMGPMGRPPGPTGAPGPQG

N-原肽

螺旋结构域

FQGNPGEPEGAAGPMGR GPPGPPGKPGDDGETGKPGKSGERGPPGPQGARGFPCTPGLPGVKHGRYPGLDGAAG  
EAGAPGAKGESGSPGENGSPGPMGRGLPGERGRPGSGAAGARGNDGLPGAPGPPVGPAGAPGFPAGPSKGEAGPTGARGPE  
QAQGRGESGTPGSPGAPAGPNPGTDGIPGAKGSAGAPGIAGAPGFPGRPPGPQGATGPLGPKGQTGEPIAGFKGEQGPKE  
TGAPGQAGAPGAGEEGKRGARGEPGAAGPVGPPGERGAPGNRGFPQDGLAGPKGAPGERCPAGLAGPKGATGDPGRPGEPGLPG  
ARGLTGRPGDAGPQGVGPTGAPGEDGRPGPPGPQARGQPGVMGFPGKANGEPGKAGEKGLPGAPGLRGLPGKDGETGAAGPP  
GPAGPVGERGEQAGPSPGFQGLGPPGPPGESGKPGDQGVPEAGAPLVGPRGERGFPGERGSPGAQGLQGRGLPGTPTDGP  
KGATGPAGPNAQGPGLQGMPPGERGAAGIAGLKGRDGVGEKPEGAPGKDARGLTGPIGPPGAPNGEKGESGPPGSGAAG  
ARGAPGERGEPGAPGAGFAGPPGADGQPGAKGEQGEQKGDAGAPGQPSGAPGQGPPTGVTGPKARGAQGPATGFPAA  
GRVGPPGPNNGPPGPPGSAGKDGPKGVRGDAGPPGRAGDPLQGPAGPPGEGKEPGEEDGPAGPDGPPGQLAGQRGIVGLPGQ  
RGERGFPGLPGSPGEPCKQAGPSAGDRGPPGVPGLTGAGEPGREGNPGADGLPGRDGAAGVKGRGETGPVGAPGAPGAPG  
APGPVGPTGKQCDRETGAQCPMGPSGPAGRCMPGQGPGRDKGETGEAGERGLKGRGFTGLQGLGPPGPGSGDQGAAGPAGPS  
GPRGPPGPV

C-原肽

GPSGKDGSGMPGPIGPPGPRGRS GEPGAPGPPGNPGPPGPPGPP GTGIDMSAFAGLQTE KGPDPTRYMRA DEA

C-原肽

AGGLRQHIDVEVDATLKSNNQIESIRSPEGSKKNPARTCRDIKLCHPEWKSQDYWIDPNQGCTLDAIKVFCNMETGET  
CVYPTPSSIPRKNWWTSTKDKKHVWFAETINGGFHFSYGDENLSPNTASIQMTFLRLLSTEGSQNVTYHCKNSIAYMDEETGNLK  
KAILIQGSNDVEIRAEGNSRFTYSVLEDGCTKHTGKWGKTVEYRLQKTSRLSTVDTPMDIGGADQEFQVDIGPVCHL

鸡 II 型胶原结构

图 3

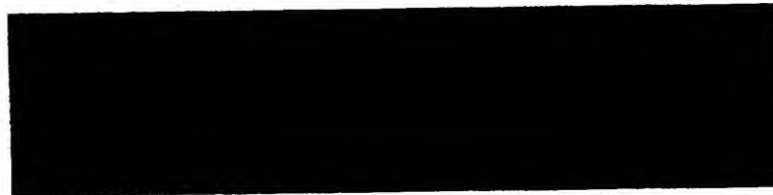
M 1 2 3 4 5 6 7 8 9 10 11 12 13 14 M



17 日龄鸡胚 3' UTR PCR 结果

- 1、心脏, 2、肝脏, 3、眼球玻璃体, 4、角膜, 5、皮肤, 6、胸腺, 7、胸肌, 8、胸骨  
9、小肠, 10、关节软骨, 11、脾脏, 12、半月板, 13、颅骨, 14、睾丸

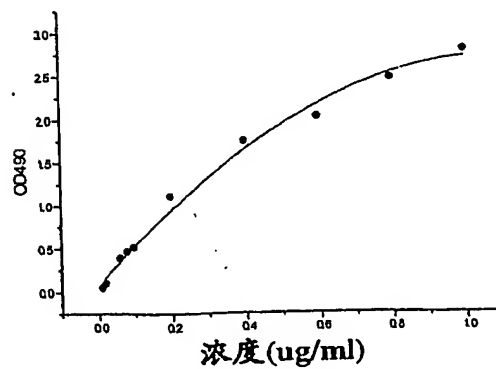
图 4



成鸡 3' UTR PCR 结果

- 1、心脏, 2、肝脏, 3、眼球玻璃体, 4、角膜, 5、皮肤, 6、胸腺, 7、胸肌, 8、胸骨  
9、小肠, 10、关节软骨, 11、脾脏, 12、半月板, 13、颅骨, 14、睾丸

图 5



CCII 浓度与吸光度标准曲线

图 6

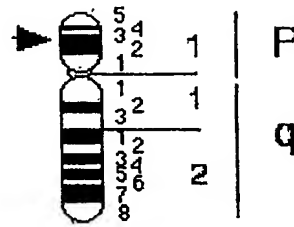
1 2 3 4 5 6 7 8 M



CCII 外显子 2 剪切分析

1、肝脏，2、皮肤，3、眼球玻璃体，4、小肠，5、胸肌，6、角膜，  
7、关节软骨 8、胸骨 9、M

图 7



染色体分带

图 8



鸡染色体中期分裂相



图 9

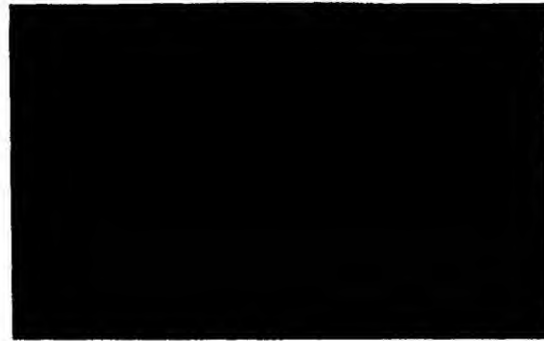


图 9 ISH 杂交结果

图 10

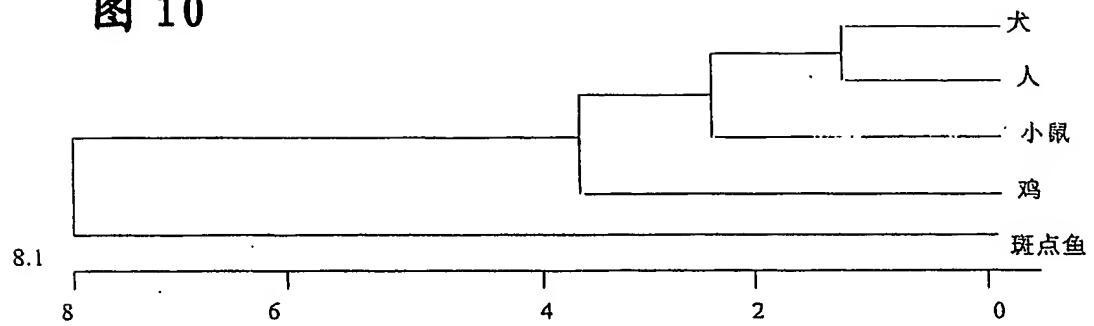


图 10: 人、犬、鼠、鸡、斑点鱼 CCII 同源性比较建立进化树